# The Absolute-Value Estimate for Symmetric Multilinear Forms*

William C. Waterhouse
*Department of Mathematics*
*The Pennsylvania State University*
*University Park, Pennsylvania 16802*

Submitted by Richard A. Brualdi


## ABSTRACT

Let $T(v_1,\dots,v_n)$ be a symmetric multilinear function on vectors $v_1,\dots,v_n$ in some $\mathbb{R}^r$. Let $M=\max_{\|v\|=1}|T(v,\dots,v)|$. Then $|T(v_1,\dots,v_n)|\leqslant\|v_1\|\cdots\|v_n\|M$. Equality (for nonzero $v_i$) can occur only in very special cases when the $v_i$ are not all parallel, and never if they span more than a plane. The inequality is known, but the proof here is new, as are the precise conditions for equality.


## 1. THE THEOREMS

Let $T(v_1,\dots,v_n)$ be a symmetric multilinear function on vectors $v_1,\dots,v_n$ in some $\mathbb{R}^r$, and let $F(v)=T(v,\dots,v)$ be the corresponding homogeneous polynomial function of degree $n$ on $\mathbb{R}^r$. It is well known that $T$ is uniquely determined by $F$. But there is also a relation between their values:

THEOREM 1. *Let* $M=\max_{\|v\|=1}|F(v)|$. *Then*

$$|T(v_1,\dots,v_n)|\leqslant\|v_1\|\cdots\|v_n\|M.$$

Equivalently, this theorem says that the maximum absolute value of $T$ on unit vectors can be attained with all the $v_i$ equal. The theorem is by no

*This work was supported in part by the U.S. National Science Foundation, Grant DMS8701690.




655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

means new (see Section 2), but I think my proof will be shorter and more "algebraic" than any currently available. I shall also go on to determine precisely when equality in Theorem 1 can occur with the $v_i$ not all parallel. Essentially the only example is the one described in the following definition, which seems to be easiest to understand if we state it using complex numbers:

DEFINITION. An *equality plane* for $T$ (or for $F$) is a two-dimensional subspace of $\mathbb{R}^r$ having an orthonormal basis $e$, $f$ such that $T(s_1e + t_1f, \dots, s_ne + t_nf) \equiv M \operatorname{Re}[(s_1 + it_1) \cdots (s_n + it_n)]$ or (equivalently) $F(se + tf) \equiv M \operatorname{Re}[(s + it)^n]$.

Note carefully that $M$ here is still meant to be the maximum of $F$ on the whole unit sphere in $\mathbb{R}^r$, not just on the unit circle in the plane of $e$ and $f$.

THEOREM 2. *If equality in Theorem 1 occurs for nonzero vectors $v_i$ spanning a plane, then their span is an equality plane.*

THEOREM 3. *Equality in Theorem 1 never occurs for nonzero vectors spanning a subspace of dimension more than 2.*

A closely related statement about the gradient $\nabla F$ also needs to be recorded:

COROLLARY 4.

(a) *Let $F$ be a homogeneous polynomial of degree $n$. Let $M$ be its maximum absolute value on the unit sphere. Then for all vectors $w$ we have $\|\nabla F(w)\| \leqslant nM\|w\|^{n-1}$.*

(b) *If equality occurs in* (a) *with $\nabla F(w)$ not parallel to $w$, then $w$ and $\nabla F(w)$ span an equality plane for $F$.*

(c) *Conversely, whenever $w$ lies in an equality plane for $F$, then $\nabla F(w)$ lies in that plane and $\|\nabla F(w)\| = nM\|w\|^{n-1}$.*

It is easy to prove Corollary 4 from Theorems 1 and 2 once we recall the classical idea of polarization (or use direct computation on monomials) to see that $\nabla F(w) \cdot v = nT(v, w, \dots, w)$. Choosing $v$ to be a unit vector in the direction of $\nabla F(w)$, we get (a) from Theorem 1. If this direction is not parallel to $w$, then Theorem 2 gives us (b). Part (c) now is a straightforward computation. Indeed, for $w = se + tf$ we have $\partial F(se + tf)/\partial s = M \operatorname{Re}[n(s + it)^{n-1}]$ and $\partial F(se + tf)/\partial t = -M \operatorname{Im}[n(s + it)^{n-1}]$. These two components of $\nabla F(se+tf)$ make a contribution of size $nM\|w\|^{n-1}$ to the norm. By (a) they must give us the whole gradient, which thus lies in the plane of $e$ and $f$.

## 2. A BIT OF HISTORY

So far as I can discover, Theorem 1 was first proved in 1935 by van der Corput and Schaake [4]. They began by proving an inequality on trigonometric polynomials (see Section 4), deducing Corollary 4(a), and then obtaining Theorem 1 by a simple induction (compare the last paragraph in Section 3). But soon they found [5] that Corollary 4(a) had been proved in the same way back in 1927 by O. D. Kellogg [7]. Thus probably Kellogg should get some of the credit for Theorem 1.

For some years after that, however, apportioning the credit would have seemed pointless; for Theorem 1 disappeared from the collective mathematical memory. It turned up again in a 1974 engineering paper by D. Ho [6], clearly a fresh rediscovery. Unfortunately, the argument given there (using Lagrange multipliers) is erroneous: Ho tried to show that equality cannot occur for unit vectors unless each component of each $v_i$ is $\pm$ the corresponding component of $v_1$, and that assertion is false on equality planes. A valid proof along the same lines was given just recently by R. A. Bousfield [3]. While mistaken in the belief that the theorem was new, Bousfield did introduce a new idea in the proof, combining Theorem 1 with a weak form of Theorem 2 to get a proof by induction. My proof will use that same idea, but the more detailed description of equality planes in Theorem 2 here allows me to bypass the use of Lagrange multipliers.

## 3. PROOF OF THEOREMS 1 AND 2

We first need a few observations about equality planes. Since we shall meet these planes in rotated coordinates, let us define

$$T_n^\lambda[(s_1,t_1),\dots,(s_n,t_n)] = \operatorname{Re}[\lambda(s_1+it_1)\cdots(s_n+it_n)],$$

where $\lambda$ is a complex number of absolute value 1. Clearly $T_n^\lambda$ does attain its maximum absolute value on unit vectors with all $s_j+it_j$ equal (namely, equal to $\lambda^{-1/n}$). The following rudimentary computation then is as close as we need to come to "hard" analysis:

LEMMA 5. *Let $c \neq 0$ be real. Let $\lambda$ be a complex number of absolute value 1. Let $n \geqslant 2$, and assume $\lambda \neq \pm 1$ if $n = 2$. Then there are complex numbers $z_j = s_j + it_j$ of absolute value 1 for which* $\mathrm{Re}[\lambda z_1 \cdots z_n] + ct_1 \cdots t_n$ *has absolute value bigger than* 1.

*Proof.* We prove this just by listing values of $z_j$ for which it holds. If $n > 2$, take any $z_1$ with $\lambda z_1 \neq 1$ and $ct_1 > 0$ and let $z_2 = \cdots = z_n$ be $(n-1)$st roots of $1/(\lambda z_1)$ with positive imaginary parts. If $n = 2$ and $c > 0$, let $z_1 = z_2$ be square roots of $1/\lambda$ with positive imaginary parts. If $n = 2$ and $c < 0$, let $z_1$ be a square root of $-1/\lambda$ with positive imaginary part, and let $z_2 = -z_1$. ∎

LEMMA 6. *When the space is $\mathbb{R}^2$, the maximum of $T$ on unit vectors is indeed attained with all $v_i$ equal. If it also occurs at unit vectors that span $\mathbb{R}^2$, then this plane is an equality plane.*

*Proof.* We use induction on $n$, there being nothing to prove when $n = 1$. Let $M_1$ be $\max|T(v_1,\ldots,v_n)|$ for unit vectors in the plane. We have nothing to prove unless that maximum is attained for $v_1,\ldots,v_n$ that span the plane. By symmetry of $T$, we can then assume that $v_1$ and $v_2$ are independent. Rotation in the plane preserves the hypotheses and the conclusion, so we may suppose that $v_n = (1,0)$. Let $T_* = T(-,\ldots,-,v_n)$. The maximum absolute value of $T_*$ on unit vectors is again equal to $M_1$. If $n > 2$, it attains that maximum on a set spanning the plane, and by induction we know that $T_* = M_1 T_{n-1}^{\lambda}$ for some $\lambda$. If $n = 2$, every linear $T_*$ has such an expansion; and since the maximum is attained with $v_1$ not parallel to $v_2 = (1,0)$, we know that in this case $\lambda \neq \pm 1$.

Now if $U[(s_1,t_1),\ldots,(s_n,t_n)]$ is a multilinear form on $\mathbb{R}^2$ vanishing identically when we set $(s_n,t_n) = (1,0)$, it cannot have any terms involving $s_n$. If it is also symmetric, its terms cannot involve any $s_j$, and we have $U$ a multiple of $t_1 \cdots t_n$. Applying this reasoning to $U = T - M_1 T_n^{\lambda}$, we see that $T = M_1(T_n^{\lambda} + ct_1 \cdots t_n)$ for some constant $c$. Since $M_1$ is the maximum absolute value of $T$ on unit vectors, Lemma 5 shows that $c = 0$, and thus we have an equality plane. We observed earlier that $M_1 T_n^{\lambda}$ then does attain its maximum on unit vectors with all vectors equal. ∎

This lemma already proves Theorem 2, since we can just restrict $T$ and $F$ to the plane spanned by the $v_i$ there. For Theorem 1, we use induction on $n$. Scaling the $v_i$, we can assume they are all unit vectors. The function $T(-,\dots,-,v_n)$ is symmetric multilinear in $n-1$ vectors, and so by induction there is some unit vector $w$ with $|T(v_1,\dots,v_{n-1},v_n)| \leqslant |T(w,\dots,w,v_n)|$. Now Lemma 6 applied to a plane containing $v_n$ and $w$ shows that there is some $u$ in that plane with $|T(w,\dots,w,v_n)| \leqslant |T(u,\dots,u,u)| = |F(u)| \leqslant M_1$.

## 4. A BIT MORE HISTORY

As we observed, Theorem 1 is roughly equivalent to Corollary 4(a). To prove that result, Kellogg as well as van der Corput and Schaake first proved a theorem that extends a famous trigonometric inequality due to Bernstein. The fundamental idea behind such inequalities has been clarified since the time they wrote, and so it may be interesting to sketch how one might give a modern proof of Corollary 4(a) along their lines.

Looking at the plane containing $w$ and $\nabla F(w)$, we see that it suffices to consider homogeneous polynomials $F(x,y)$ in two variables. By homogeneity, it suffices to prove the inequality when $(x,y) = (\cos\vartheta, \sin\vartheta)$ is on the unit circle. Let $g(\vartheta) = F(\cos\vartheta, \sin\vartheta)$. The normal component of $\nabla F$ along the circle is $x\,\partial F/\partial x + y\,\partial F/\partial y$, which by Euler's equality is $nF = ng(\vartheta)$. The tangential derivative, on the other hand, is clearly $g'(\vartheta)$. Thus the inequality to be proved says that $[|g'|^2 + n^2|g|^2]^{1/2} \leqslant n \max_\vartheta |g(\vartheta)|$. The fundamental idea, nicely explained from scratch at the start of Chapter 11 of Boas [2], is an identity satisfied not only by our trigonometric polynomial $g$ but also by any entire function $h$ having exponential growth of order $\leqslant n$: if $a$ and $\vartheta$ are real, then

$$(\sin a)h'(\vartheta) - n(\cos a)h(\vartheta)$$
$$= n \sum_{k=-\infty}^{\infty} (-1)^{k-1} \left( \frac{\sin a}{a - k\pi} \right)^2 h\left( \vartheta + \frac{k\pi - a}{n} \right).$$

This is proved by a Fourier expansion. Clearly for real-valued $h(\vartheta)$ we can choose $a$ to make the left side equal to $(|h'|^2 + n^2|h|^2)^{1/2}$, while the right-hand side is easily seen to be at most $n \max_\vartheta |h(\vartheta)|$. Our inequality in this form is Theorem 11.4.8 in Boas [2, p. 215]. The extremal nature of the functions $g(\vartheta) = M\cos(n\vartheta + c)$, corresponding to our equality planes, is also on record in this context (Achieser [1, p. 141]).

## 5. CUBIC FORMS IN THREE VARIABLES

The rest of the paper is devoted to Theorem 3, which I believe is completely new. Most of the work can be done on homogeneous cubic polynomials on $\mathbb{R}^3$, and the first such result might have some independent interest.

PROPOSITION 7. *Let $F(x,y,z)$ be a homogeneous cubic in three variables. If $F$ has more than one equality plane, then a suitable orthogonal change of variables will reduce $F$ to a constant times $x^3 - 3xy^2 - 3xz^2$.*

*Proof.* Note that orthogonal changes of variable do take equality planes to equality planes. Suppose now that $F$ has two equality planes. We may suppose that the maximum absolute value of $F$ on unit vectors is 1. By rotation, we may suppose that one of the equality planes is $z=0$. The two equality planes meet in a line; and if we take a unit vector $v$ on that line, we see by Corollary 4(c) that $\nabla F(v)$ lies in both planes and thus is parallel to $v$. Hence also $v$ gives the maximum absolute value of $F$. Replacing $v$ by $-v$ if necessary and rotating around the $z$-axis, we can suppose that the two planes include $(1,0,0)$ and that $F(x,y,0) = \mathrm{Re}[(x+iy)^3] = x^3 - 3xy^2$. Since $\nabla F(x,y,0)$ lies in the equality plane $z=0$, we see that $F$ cannot contain terms in $x^2z$ or $y^2z$. Hence we have

$$F = x^3 - 3xy^2 + axz^2 + byz^2 + cz^3$$

for constants $a$, $b$, $c$.

As the second equality plane contains $(1,0,0)$, it has an orthonormal basis consisting of $(1,0,0)$ and $(0,p,q)$ for some $p$, $q$ with $q \neq 0$ and $p^2+q^2=1$. Computation gives

$$F[s(1,0,0)+t(0,p,q)] = s^3 - 3st^2p^2 + ast^2q^2 + bpq^2t^3 + cq^3t^3.$$

This must equal $\mathrm{Re}[\lambda(s+it)^3]$ for some $\lambda$, and the coefficient of $s^3$ shows us that $\lambda = 1$. The coefficient of $st^2$ then must be $-3$, and this implies that $a = -3$ (since $q \neq 0$ and $p^2+q^2=1$). Thus

$$F = x^3 - 3x(y^2+z^2) + byz^2 + cz^3.$$

The values of this $F$ now on unit vectors, where $x^2+y^2+z^2=1$, are

$4x^3 - 3x + (by + cz)z^2$. If either $b$ or $c$ is nonzero, we can choose $y$ and $z$ with $y^2 + z^2 = \frac{3}{4}$ and $(by + cz)z^2 > 0$; then $(-\frac{1}{2}, y, z)$ would give a unit vector where $|F|$ is bigger than 1. As this is a contradiction, we conclude that $b = c = 0$. ■

**Remarks.**

(1) In the exceptional case, it is easy to check that every plane containing $(1,0,0)$ is an equality plane.

(2) It is possible for a cubic form in four variables to have exactly two equality planes. For example, the form

$$F = x^3 - 3xy^2 + w^3 - 3wz^2$$

has maximum absolute value 1 on unit vectors and has just the two equality planes given by $x = y = 0$ and $w = z = 0$.

**Proposition 8.** *Let $T(v_1, v_2, v_3)$ be the symmetric trilinear form on $\mathbb{R}^3$ corresponding to $x^3 - 3x(y^2 + z^2)$. If the $v_i$ are unit vectors with $|T(v_1, v_2, v_3)| = 1$, then the $v_i$ all lie in an equality plane.*

*Proof.* Explicitly, we have

$$T = x_1x_2x_3 - (x_1y_2y_3 + y_1x_2y_3 + y_1y_2x_3) - (x_1z_2z_3 + z_1x_2z_3 + z_1z_2x_3).$$

We know the $v_i$ giving absolute value 1 are giving the maximum absolute value on unit vectors, and hence the coefficients of the linear function $T(-, v_2, v_3)$ must form a vector in the same direction as $v_1 = (x_1, y_1, z_1)$. (This is a sort of rudimentary Lagrange-multiplier argument.) Looking at the first two coefficients, we see we have

$$x_1(-x_2y_3 - y_2x_3) = y_1(x_2x_3 - y_2y_3 - z_2z_3).$$

Rearranging the terms, we get

$$y_1z_2z_3 = (y_1x_2x_3 + x_1y_2x_3 + x_1x_2y_3) - y_1y_2y_3. \qquad (*)$$

Similar reasoning applies when we let the other $v_i$ vary, and so we get similar equations for $z_1y_2z_3$ and $z_1z_2y_3$. But the right-hand side of ($*$) is symmetric, and thus we can conclude that

$$y_1z_2z_3 = z_1y_2z_3 = z_1z_2y_3.$$

Suppose first that all the $z_i$ are nonzero. Then $y_1/z_1 = y_2/z_2 = y_3/z_3 = s$ for some $s$. Thus all three $v_i$ lie in the equality plane $y = sz$. Suppose next that $z_1$ (say) is zero but $z_2$ and $z_3$ are nonzero. The equations then give $y_1 = 0$, so (after a change of sign) $v_1 = (1,0,0)$. As a function of $v_2$ and $v_3$, then, $T$ is $x_2x_3 - y_2y_3 - z_2z_3$. Again, since this is at an extreme as a function of the unit vector $v_2$, that vector must equal $\pm(x_3, -y_3, -z_3)$. It follows that $(x_3, y_3, z_3)$ is in the plane spanned by $(1,0,0)$ and $(x_2, y_2, z_2)$, and again the vectors are in an equality plane.

Finally, suppose $z_1 = z_2 = 0$. Then $T$ is given by the formula $(x_1x_2 - y_1y_2)x_3 - (x_1y_2 + y_1x_2)y_3$. If $z_3$ were nonzero, we could change $v_3$ to another unit vector with $z_3 = 0$ and $x_3$ and $y_3$ scaled by a factor bigger than 1; this would make the value of $T$ larger. Hence we must have $z_3 = 0$, and our vectors lie in the plane $z = 0$. ∎

## 6. PROOF OF THEOREM 3

Suppose we have equality in Theorem 1 with $\dim(\Sigma\mathbb{R}v_i) \geqslant 3$. By the symmetry of $T$, we may assume that the first three vectors $v_1$, $v_2$, $v_3$ span a space of dimension 3. Let $T_* = T(-,-,-,v_4,\ldots,v_n)$. This is a symmetric trilinear form on the 3-space spanned by $v_1, v_2, v_3$, and its maximum absolute value on unit vectors occurs at that triple of independent vectors. Looking at $T_*(v_1,-,-)$, we know by Theorem 1 that its maximum must be attained also at some $T_*(v_1, w_1, w_1)$ with $w_1$ in the span of $v_2$ and $v_3$. Here $v_1$ and $w_1$ must be independent, so by Theorem 2 we know that $v_1$ and $w_1$ span an equality plane for $T_*$. Similarly $v_2$ and some $w_2$ span an equality plane for $T_*$, as do $v_3$ and some $w_3$. Since the $v_i$ are independent, there must be at least two equality planes for $T_*$. Hence $T_*$ is of the special type given in Proposition 7. But Proposition 8 shows that such a form does not reach a maximum for independent $v_1$, $v_2$, $v_3$. Thus our equality cannot occur at all. ∎

## REFERENCES

1 N. I. Achieser, *Theory of Approximation* (transl. by C. J. Hyman), Ungar, New York, 1956.
2 R. P. Boas, *Entire Functions*, Pure Appl. Math. 5, Academic, New York, 1954.

3 R. A. Bousfield, A theorem for symmetric $n$-linear forms, *Linear Algebra Appl.* 105:183–193 (1988).
4 J. G. van der Corput and G. Schaake, Ungleichungen für Polynome und trigonometrische Polynome, *Compositio Math.* 2:321–361 (1935).
5 ____, Berichtigung, *Compositio Math.* 3:120 (1936).
6 D. Ho, Buckling loads of nonlinear systems with multiple eigenvalues, *Internat. J. Solids and Structures* 10:1315–1330 (1974).
7 O. D. Kellogg, On bounded polynomials in several variables, *Math. Z.* 27:55–64 (1927).

*Received 9 September 1988; final manuscript accepted 7 December 1988*